Wirtschaftsinformatik 2021 Proceedings

Track 18: Future of Digital Markets and Platforms

Data-driven Competitive Advantages in Digital Markets: An Overview of Data Value and Facilitating Factors

Victoria Fast
Universität Passau

Daniel Schnurr
Universität Passau

Michael Wohlfarth
Universität Passau

Data-driven Competitive Advantages in Digital Markets: An Overview of Data Value and Facilitating Factors

Victoria Fast[1], Daniel Schnurr[1], and Michael Wohlfarth[1]

[1] University of Passau, School of Business, Economics and Information Systems, Passau, Germany
{victoria.fast, daniel.schnurr, michael.wohlfarth}@uni-passau.de

Abstract. Recent high-profile merger and antitrust cases as well as policy debates worldwide have focused on the relationship between access to (big) data and firms' competitive advantages in digital markets. These discussions have brought forward numerous conceptual arguments for and against the conjecture that market power may be derived from a firm's access to big data. Based on a review of the economic, information systems and management literature, this paper presents an overview of the aggregate empirical evidence on the business value and economic benefits that firms can indeed create from big data in the Internet economy. Moreover, six facilitating factors for data-driven market power are proposed that enable a firm to establish a sustained competitive advantage based on the economic benefits from data. Finally, we point to policy measures which may address competitive concerns in data-driven digital markets and highlight opportunities for future information systems policy research.

Keywords: big data, data-driven business models, competition in digital markets, market power, regulation, policy, online platforms, Internet economy

1 Introduction

Digital technology creates a wealth of data, and has therefore given "rise to a new economy" [1]. In this vein, data itself is recognized to have economic value as a resource. The European Commission, for example, has vowed to build the European Data Economy, acknowledging that "access to data spurs marketplace efficiency and innovation" [2, p. 10]. At the same time, the Internet economy, which has been at the forefront of digitization, has born a small number of "superstar firms" [3], whose success is often attributed to their superior access to data as well as their skills in exploiting and monetizing this data. Whether the success of these firms translates into a long-term economic benefit for all stakeholders is currently controversially debated. Sparked by several high-profile antitrust cases (see, e.g., the cases *European Commission v. Google* and *Bundeskartellamt v. Facebook*) as well as ongoing investigations scrutinizing the most popular Internet content and service providers (CSPs), it is frequently questioned, whether the data-driven economic benefits achieved by these firms can be imitated or leapfrogged by competitors and market entrants, or

16th International Conference on Wirtschaftsinformatik,
March 2021, Essen, Germany

whether sustained competitive advantages may protect market power and enable anti-competitive behavior in the long run. The answer to this central question has important ramifications for the competitiveness and innovation in digital markets.

Thus, the primary goal of our study is to inform policy makers about the effects of big data as a competitive resource in digital markets and guide policy interventions that aim to address competitive concerns about dominant data-rich firms, especially in the Internet economy. Although the information systems literature has analyzed the economic and strategic value of data in numerous use cases and with regard to various performance metrics, the relationship between data as a resource and a firm's ensuing economic benefits has not been investigated based on the available, consolidated empirical evidence. Moreover, the management and economics literature have emphasized that additional factors need to be considered when assessing whether short-term economic benefits will translate into sustained competitive advantages. In this article, we address these issues by reviewing the empirical evidence on data-driven economic benefits in the Internet economy and characterizing the conditions that make it likely that data-rich firms can obtain dominant market positions.

We contribute to the academic literature and the policy debate by categorizing the various economic benefits that firms can achieve from the collection and use of data resources. To this end, we focus on the personalization of services and recommendations as well as targeted advertising as the main big data use cases in the Internet economy. By highlighting moderating factors that have been found to significantly influence economic benefits in these use cases, we present more nuanced insights on how big data affects the economic performance of firms. This contributes to an improved understanding of the actual business value of (big) data. In addition, we identify six facilitating factors which enable a firm to establish a competitive advantage based on big data resources. In particular, we suggest that (i) exclusive access to data, (ii) exploitative access to data, (iii) economies of scale in data analytics, (iv) network effects and platform business models, (v) data-induced switching costs, and (vi) digital services ecosystems and economies of scope can protect a firm's data-driven competitive advantage from imitation by its competitors. We conclude that only in use cases where data creates significant economic benefits and where such benefits are additionaly protected by the identified facilitating factors, it is likely that firms can sustain a data-driven competitive advantage and establish a dominant market position.

From a policy perspective, these factors provide suitable starting points for designing remedies to address competitive concerns about data-driven market power. To this end, we point to a set of escalating policy measures for protecting competition in digital markets. We emphasize that future research on the interplay of these measures and their respective effectiveness is needed. In this vein, our review also serves as a basis for future information systems policy research that aims to tackle the theoretical, empirical and design questions that come with the need to govern data-driven digital markets.

2 Methodology

Our study comprises two main parts. First, we conduct a literature review [4-6] on the empirical evidence of economic benefits from big data. Following Schwarz et al. [5], the main purpose of this review is to summarize the joint empirical evidence on different categories of economic benefits that have been identified in individual studies. Given our primary goal to inform the current policy discourse on data-driven competitive advantages in the Internet economy, we limit the scope of our literature review to studies that have investigated use cases and business models where big data has been used for data-driven quality improvements, service personalization, recommendations and targeted advertising. Thus, we focus on use cases of (i) big data and predictive data analytics [7] (rather than descriptive or prescriptive data analytics [8]) (ii) in core business and operational functions [9] and (iii) scenarios where big data regularly consists of user data. In consequence, this leaves out studies that have investigated firms' use of big data for strategic decision making or enhanced decision support in broader industry contexts, which have been studied in-depth by the business value and business analytics literature [see, e.g., 10-12]. Instead of striving for general exhaustiveness [6], limiting the scope and narrowing our focus to *big user data* allows us to (i) synthesize and consolidate the relevant empirical evidence for use cases that are currently the focal point of the ongoing policy debate and (ii) derive more nuanced insights by detailing the most important influencing factors that moderate the effectiveness of big data in generating economic benefits in these use cases.

As our research question transcends individual disciplines and has been studied in the information systems, marketing, management and economics literature, we do not limit our analysis to a set of specific journals or disciplines. Instead, we have initiated our literature search from a set of well-known and influential articles on specific big data use cases published in the top journals of these fields and then conducted extensive backward and forward searches [4,6]. This was complemented by iterative keyword searches for the categories of economic benefits identified in the literature [6].

In the second part of this article, we build on the review in the first part and draw on the resource-based view of the firm [13] to derive a framework of facilitating factors that allow firms to build a sustained competitive advantage based on the economic benefits from data. To operationalize the concept of a competitive advantage and retrieve the relevant literature, we focus on the key requirements of a firm’s valuable and “inimitable resources and capabilities” [12, p. 357] that are necessary to establish a sustained competitive advantage [14-16]. As the literature on competitive advantages from big data is much less mature, we rely predominantly on studies that employ economic theory, conceptual analyses and empirical investigations of specific use cases to derive our proposition on key facilitating factors. In consequence, we do not claim the proposed set of facilitating factors to be definite or exhaustive. Instead, we synthesize the combined insights of the ongoing interdisciplinary research efforts. Thus, our review also reveals opportunities for future theoretical and empirical work.

3 Data Value Creation: Economic Benefits from Big Data Use

The progress and diffusion of information technology together with the continuing shift of consumers’ activities to digital (online) environments has led to a rapid growth of data that is being collected, stored and analyzed [17]. Such *big data* consists of large, heterogeneous, unstructured or semi-structured and generic-purpose data sets [18]. In the Internet economy, CSPs collect data mainly in the form of transaction logs, which are created when consumers’ behavior and actions are tracked and recorded (*big user* *data*) [19-21]. In practice, big data is often characterized based on the 3-V-Model referring to volume, variety and velocity [22]. The latter highlights that steady updateability of this data is a crucial aspect next to the sheer volume and heterogeneity, when collecting and processing big data [18]. This is emphasized by the metaphor of data flows instead of data stocks [9]. Moreover, collected web data is typically sparse, implying that it "consists of individually rare, but collectively frequent events" [23, p. 9].

Big user data collected by online firms often contains information, which can – either directly or combined with additional data – be used to identify individuals [24]. Therefore, big data is frequently also *personal data*. This data may either be actively provided by users themselves, e.g., in the case of online purchases or the creation of a user account in social networks [25], or may be collected by tracking users' online activities, e.g., by logging their browsing, search and purchasing behavior. Thus, the online environment allows for "fast, easy and unobtrusive collection of detailed information on individual activities" [26, p. 35]. Collected data may further be enriched by inference of new information through linking and analyzing existing data sets [25]. Firms may not necessarily rely (exclusively) on their own data collection, but also purchase data from external sources or adopt marketing services from customer data intermediaries [27-28].

3.1 Data-driven Quality Improvements and Service Personalization

The collection and processing of data enables CSPs to continuously improve the quality of their offerings. The analysis of individual clickstream data (possibly enriched by additional data sets) allows CSPs to derive general insights about users' browsing behavior and thus provides opportunities to adapt the design and navigation of user-facing interfaces according to their primary business goals [26], [29]. For example, advertising-financed websites want to maximize website visits and duration of visits, whereas e-commerce websites benefit from terminating website visits early with a purchase (see [26] for an overview of clickstream data analysis in e-commerce). In addition, CSPs can exploit clickstream data to improve the presentation of content as well as navigation elements by running usability studies and field experiments and simultaneously tracking changes in consumer behavior [30]. In the context of search engines, tracking data which records user behavior (specifically, search queries and clicking behavior on the search engine's results page) enables providers to refine their search algorithm and to align it more appropriately with users' needs. By doing so, the perceived quality of search results can be improved [25]. Yao and Mela show that search engines can increase their revenues from sponsored search advertising by designing the user and search interface based on the analysis of log files that track user behavior [31]. Moreover, the continuous analysis of users' interactions enables established CSPs to identify shifting demand patterns. In this spirit, Du and Kamakura develop a "quantitative trendspotting" approach, which they apply to online keyword search data to uncover consumer demand trends in the retail automotive industry [32]. Moreover, by continuously running experiments to test the effects of incremental updates, established CSPs can identify relevant and valuable new offerings, thus increasing the rate of innovation through the collection and analysis of data [33].

Next to these general quality improvements, the collection and processing of data is the basis for the *personalization of content and services* [34]. Thereby, personalization is generally understood as the delivery of “the right content to the right person at the right time to maximize immediate and future business opportunities” [35, p. 867]. To this end, internal tracking data and data from external resources are collected, aggregated and processed to create individual user profiles, which approximate users’ true interests and preferences [30]. Moe demonstrates that online shops can categorize user visits according to shoppers’ motivation and predict individual purchasing likelihoods based on observed clickstream data [36]. Today, online retailers suggest delivery and payment options based on previous orders [37]. Furthermore, search engines use individual search histories and the current location of users to adjust the ranking of search results, while websites and especially social networks frequently adapt displayed content to the interests of users [33]. The ability to personalize media content on a large scale has driven the success of online social media services, thereby giving rise to *social big data* [38]. Processing and analyzing social big data can, e.g., be used to target advertisements (see Section 3.3) or provide user-experience-based visualization, which can generate new insights for users and thus increase the perceived value of the service [38]. Moreover, collaborative processing of unstructured social data enabled by IT systems has been found to foster exploratory innovation [39].

It is widely recognized that personalization can positively affect *users’ satisfaction* and retention as well as cross-selling opportunities [40-41]. Because search and transaction costs for users are generally low in online markets [42], competition has often been claimed to be “only a click away" [43]. In consequence, personalization is a particularly important mechanism to foster *customer loyalty*. Moreover, personalized purchase processes and offers can increase users’ switching costs and thus impede customer poaching by competitors [40]. In a field experiment, Benlian shows that giving users the possibility to personalize the content and design of a website may influence users’ willingness to stick with a website and also their willingness to pay [44]. Based on a sample of 422 “brick-and-click” as well as online-only retailers, Thirumalai and Sinha show that personalizing (purchase) transactions improves customer loyalty for most, but not for all of the retailers [45]. Thus, they highlight that the benefits of personalization hinge on several specific characteristics of online CSPs, e.g., product selection and variety [45].

As personalization is used to tailor content and services, it also affects the *differentiation between firms’ goods* and the relative perception by consumers. In this regard, game theory analyses show that personalization based on consumers’ purchase histories can also lower profits if online CSPs compete with each other [46]. In particular, personalization intensifies competition in the market by reducing differentiation between two competing online services, making products more homogeneous from the perspective of consumers. Anticipating this prisoner’s dilemma-like situation, each competitor has a strategic incentive to quickly increase its market share in order to prevent rivals from gaining access to users’ purchase history. However, if online CSPs attempt to do so by offering a mainstream service, this again reduces differentiation, thus leading to more competition for the market [46].

Moreover, it is well-known that collecting and processing data of users to personalize content and services can elicit *privacy concerns* [47-48] which have been found to negatively impact the likelihood of using a personalized service [37]. Yet, Chellappa and Sin show that firms can actively mitigate these negative consequences through *trust-building measures* [37]. Specifically, the reputation of an online service, e.g., the perception of its brand image, has a significant impact on whether consumers accept or reject personalization. Thus, large and established firms are likely to hold an advantage over small firms and market entrants, when collecting and using consumers' data to offer personalized content and services.

Summary 1 (Service personalization): *Collecting and processing user data enable firms to provide personalized content and services, which offer economic benefits to firms by increasing users' satisfaction, willingness to pay, switching costs and thus, loyalty (see Table 1). However, personalization can also have detrimental effects on firms due to increased privacy concerns or intensified competition in and for the market in consequence of diminished product differentiation.*

3.2 Personalized Recommendations

User profiling and the analysis of users' transaction history have fueled the growth of recommendation agents as a core functionality of today's online CSPs. Providing consumers with personalized recommendations – based on either content, collaborative or hybrid approaches [49] – simplifies users' purchase decisions or consumption choices by highlighting offers that match their interests and by reducing their search costs [50]. It is then easier for users to discover niche products and to gain access to previously unknown or new content and products [51]. By providing a greater product variety and more tailored offerings through personalized recommendations, a CSP can attract a *larger customer base* with diverse preferences [52]. In consequence, recommender systems can have a significant impact on a firm's sales of products or the content that is consumed by its users. Several studies discuss the impact of personalized recommendations on *sales diversity* with regard to the upstream product and content markets [52-54]. On the one hand, the long-tail effect may lead to an increase of the demand for niche products. On the other hand, the superstar effect may further increase sales concentration for popular blockbuster products. Moreover, recommendations are likely to expand demand due to *cross-selling of complementary goods* that provide additional value to consumers. Recommender systems may also increase consumption simply by attracting consumers without concrete purchase intentions to buy recommended products [55]. Even in cases where recommendations lead to demand substitution rather than an overall demand expansion, they can *increase profits* for a CSP. Using a simulation calibrated with real sales and experimental data, Hinz and Eckert show that a content provider's profit increases due to personalized recommendations if niche products are associated with higher profit margins [56].

Experimental studies have investigated *behavioral effects* and show that recommendations and displayed ratings can increase users' preferences [57] and their willingness to pay [58]. Thus, personalized recommendations can directly increase revenues and profitability [58] or create indirect benefits, e.g., when firms use recommendations to manage inventory [57]. Game theory analyses of recommendations and competition show that an increase of a recommender system's effectiveness has ramifications for both personalizing and non-personalizing firms. Improvements in recommender systems can increase the differentiation between firms which sell similar products, thus leading to higher prices and profits for both firms [59]. In consequence, non-personalizing firms may also choose to freeride on improved recommendations of personalizing firms [59].

The performance of recommender systems is directly tied to the data, which is fed into the system to derive recommendations. *Availability of data* is a key prerequisite for the technical implementation of recommender systems as well as their economic success. Most systems base their recommendations on consumers' past interaction with the service [49]. To this end, any firm must overcome the well-known cold start problem, i.e., a *critical mass of user data* is required to elicit preferences and give useful product recommendations [60-61]. More generally, there is a positive relationship between more training data and more precise recommendations [19]. In particular, more comprehensive user profiles obtained from larger data sets reduce profile fragmentation and decrease firms' uncertainty regarding the implicit ratings inferred from user behavior [30], [62]. In this context, data which reveals context information (e.g., time, location or group composition of targeted users) can improve the performance of recommender systems [63]. Furthermore, more precise recommendations may propel a feedback loop, because increases in usage and the number of purchases, in turn, make more data available [64]. This may give incumbents a competitive advantage over entrants (see Section 3). In order to obtain accurate estimates of users' preferences, the underlying user profiles must be up-to-date, which requires continuous changes on the basis of new data [65-67]. Poor *data quality*, i.e., imprecise or incorrect data, could lead to ill-fitting recommendations, which negatively impact customer satisfaction [68-70].

Empirical research highlights additional factors that influence the effectiveness of personalized recommendations, such as the trade-off between *timeliness* and recommendation quality: On the one hand, recommendations should be generated in real time before users identify items by themselves as web sessions are typically very short. On the other hand, the longer the observation period, the more data about user preferences and tastes can be collected which improves the recommendation quality [71]. Thus, online services have to weigh between immediate sales and profiling when using personalized recommendations [72].

Moreover, *trust* influences the effectiveness of recommendations as it improves perceived usefulness and the intention to adopt recommendation agents [73]. To this end, firms can use *transparency measures* and *explanations* about recommendation agents to positively influence trusting beliefs of users [74]. In contrast, trust may be eroded by biased recommendations [75]. Sponsorship disclosures combined with further explanations on the recommendation agent can mitigate such detrimental effects on trust, but do not decrease distrust, caused by biased recommendations [75]. Transparency and explanations have been found to directly increase liking [76] and acceptance of recommendations [77]. However, the more recent study by Karwatzki et al. cannot confirm a positive effect of a firm's transparency regarding its data collection on users' willingness to disclose data, due to an increase of consumers' privacy concerns [78]. Thus, personalized recommendations are likely to be more effective for users that are less privacy sensitive [78-79].

Table 1. Overview of empirical studies on the effectiveness of service personalization and recommender systems

Factor	Personalization type	Reference	Context	Methodology	Effectiveness measure	Findings
Timing	Recommender systems	[71]	Book and music recommen-dations	Theoretical framework; lab and field experiments	Quality of recommendations	**(+)** of late presentation timing as more data about user preferences and tastes can be collected.
					Acceptance of recommendations	**(-)** of late presentation timing as recommendations should be generated in real time to match users' current needs.
Trust	Recommendation agents	[73]	Digital camera recommen-dations	Trust-TAM; lab experiment	Perceived usefulness; Intention to adopt recommendation agents	**(+)** of trust in agents (competence, benevolence, integrity).
	Personalized service	[37]	Different online industries	Survey	Intention to use personalized services	**(+)** of trust in vendor.
					Users' privacy concerns	**(-)** of trust-building activities.
Operating context	Transaction and decision personalization	[45]	Online retail	Econometric model; counterfactual analysis	Customer loyalty	**(+)** of transaction personalization for most of the retailers. **(+) or (-)** of decision personalization; the effect is dependent on operating characteristics of the firms (e.g., product selection and variety, scale of operations).
Type of personalization	Content and design personalization	[44]	News aggregator website	Field experiment	Website stickiness	**(+)** of content personalization; mediated by preference fit and perceived enjoyment. **(+)** of design personalization; mediated by perceived enjoyment. **(+)** of combing both personalization types on perceived enjoyment and website stickiness.
					Willingness to pay	**(+)** of content personalization; mediated by preference fit and perceived enjoyment. **(0)** of design personalization. **(0/-)** of combing both personalization types on preference fit and users' willingness to pay, compared to using only content personalization.
Transparency	Recommender systems	[78]	Event recommen-dations	Online experiment	Intention to disclose information	**(+)** of personalization itself. **(0)** of transparency features with respect to data collection and use because of increasing privacy concerns.
	Personalized service and advertising	[79]	Online services	Survey	Willingness to be profiled for personalization	**(-)** of users' valuation of information transparency.
	Recommender systems	[76]	Music recommen-dations	Survey	Liking of and confidence in recommendations	**(+)** of transparency.
	Recommender systems	[77]	Movie recommen-dations	Survey, experiment	Acceptance of recommendations	**(+)** of explanations regarding the reasoning of recommendations.
	Recommendation agents	[74]	Digital camera recommen-dations	Lab experiment	Trusting beliefs	**(+)** of explanations on users' trusting beliefs. **(+)** of *how* explanations on competence and benevolence beliefs. **(+)** of *why* explanations on benevolence beliefs. **(+)** of *trade-off* explanations on integrity beliefs.
	Recommendation agents	[75]	Digital camera recommen-dations	Lab experiments	Trust in biased recommender system	**(-)** of biased recommendations with sponsorship disclosure. **(+)** if both sponsorship disclosure and explanations for organic recommendations are provided.
					Distrust in biased recommender system	**(+)** of biased recommendations with sponsorship disclosure. **(0)** of providing sponsorship disclosure, explanations for organic recommendations or both

Note: (+) = positive effect of factor on effectiveness measure; (-) = negative effect of factor on effectiveness measure; (0) = no significant effect of factor on effectiveness measure.

Summary 2 (Personalized recommendations): *Recommender systems can increase sales or consumption provided that sufficient, accurate, current and context-dependent (personal) data is available to an online service (see Table 1). Poor data quality, biased recommendations and privacy concerns may compromise the effectiveness of recommender systems.*

3.3 Targeted Advertising

Targeting, i.e., the selection and tailoring of advertisements to the viewer on an individual basis, is of particular importance for online CSPs, as they frequently offer their services free of charge to consumers and rely on advertising as their primary revenue source [33]. To this end, the goal of targeting is to enhance *advertising effectiveness* by reaching those users who are more likely to be interested in the promoted services or products and thus are more likely to purchase them. As a higher ad effectiveness raises the net benefit for advertisers, due to less wasted ad impressions and higher conversion rates, publishers are likely to benefit from higher advertising revenues through targeting. With regard to consumers' reception, advertising has multi-faceted effects. Next to possibly being persuasive, advertising is also informative, as it lowers consumers' *search costs* [80]. From an economic perspective, it is thus generally assumed that targeting also benefits consumers through the access to more relevant offers and a reduction of annoying advertisements [81-82].

In principle, three types of targeting can be distinguished: In the case of *context-based targeting*, advertisements are displayed according to the content of a website [83]. Car manufacturers, for example, place advertisements on automotive portals [84], because they assume that a user's interest matches the content of the visited website [33]. In practice, advertising networks such as Google AdSense allow the large-scale placement of advertisements, which match the content of websites. *Segment-based targeting* splits users into homogeneous groups in order to customize advertisements based on the associated group characteristics. This segmentation is usually based on demographic characteristics or observed attributes [85], such as social networks used in the past [86], cognitive styles [87] or affinity to celebrities [88]. *Behavior-based targeting* displays advertisements based on tracking data that captures consumers' online activity, such as visited websites, performed search requests, past purchases or e-mails [83], [89]. A popular form of behavior-based targeting in e-commerce is known as *dynamic retargeting*, which displays advertisements according to a user's interest or behavior in the past (e.g., advertisements for a product that the user has recently viewed on a shopping website). Dynamic retargeting has been found to increase the conversion rate of prospective buyers [90], particularly if consumers are at an advanced stage of product search and have already specified their preferences [91]. To implement most of these targeting approaches, publishers require information about the users that will see the advertisements. Especially, behavioral targeting relies on *comprehensive user profiles* that are built based upon the collection and analysis of big data, especially transaction logs of users' clickstreams.

Recent studies provide robust empirical evidence for a general positive relationship between targeting and advertising effectiveness, whereby effectiveness has been measured by several different performance indicators [85]. Studies confirm a positive effect of targeting on *online sales* [90-91], the *purchase intention* of consumers [84], [92], *website visits* [90], [93], *click-through rates*, i.e., the number of displayed ads that users clicked on divided by the number of total impressions [40], [88], and *view-through rates*, i.e., the ratio between ad impressions that were followed by a successful sale and the total number of impressions [85].

However, the collection and analysis of data does not solely influence the effectiveness of targeted advertising through a better match between advertisements and consumers' interest. Instead, the use of data – which may be considered to be personal or inappropriate – affects *consumers' perception and acceptance* of targeted advertisements [83]. If consumers consider tracking of online activities and collection of personal data to violate their privacy, such concerns can have an indirect negative impact on the effectiveness of targeted advertising [82], [84]. Hence, stricter *privacy policies* by firms, which limit the collection and use of personal data, can, in principle, mitigate consumers' concerns and enhance the effectiveness of targeted advertising. At the same time, such constraints on the collection and processing of data may impair the targeting quality and generate less effective matches between ads and consumers. This negative effect has been confirmed by Goldfarb and Tucker in the context of the introduction of the European Privacy and Electronic Communications Directive (2002/58/EG) in 2002, which made it more difficult for advertisers to collect data for targeted advertising [94]. For publishers that were subject to the restrictions by the directive, the authors find a significant negative effect on the effectiveness of advertising, especially with respect to simple advertising banners and advertisements on websites which provide general-purpose instead of special-interest content [94].

In order to alleviate consumers' concerns and reactance, the empirical literature has identified firms' *transparency about data collection* as an important moderating factor on consumers' acceptance of targeted advertising and advertising effectiveness. In particular, greater transparency is found to have a positive effect if users otherwise realize that personal information is collected for targeting practices without their consent and thus feel more vulnerable [95]. Aguirre et al. show that targeted advertisements increase click-through rates if users are informed about data collection practices of online firms [95]. In contrast, advertising effectiveness decreases with targeting if data is collected without users' awareness. Yet, the impact of transparency also hinges critically on the *type of information* that is revealed to consumers. On the one hand, Kim et al. confirm that transparency about data collection and data use can increase ad effectiveness given that users trust the service and acceptable data collection practices are disclosed [96]. Here, data collection practices are deemed acceptable if personal data was (i) obtained within the website and not from third parties and (ii) directly provided by the users and not inferred by the firm. If, on the other hand, transparency reveals data collection practices that are deemed unacceptable, ad effectiveness decreases due to higher privacy concerns. Furthermore, Aguirre et al. find that transparency measures about data collection can also be effective if they inform consumers ex-post, i.e., at the time when a targeted advertisement is displayed [95]. For example, informational cues that notify users about data collection practices can serve as trust-building measures that offset negative feelings of vulnerability and consequently increase advertising effectiveness. In contrast, Samat et al. find that if users have negative opinions about targeting, consumers' attitudes and purchase intention suffer from disclosing information ex-post [97]. Consumers with neutral or

Table 2. Overview of empirical studies on the effectiveness of targeted advertising

Factor	Targeting method	Context	Reference	Methodology			Measure	Empirical findings on the effectiveness of targeting depending on factor influence
				Field exp.	Lab / online exp.	Survey		
Timing and placement	Retargeting (Personalized ads)	Online shopping	[85]	x			CTR	**(+)** especially for (i) high personalization in early stages and (ii) medium personalization in later stages of purchase decision.
				x	x		VTR	**(+)** if banners appear on motive congruent websites.
	Dynamic retargeting	Travel website	[91]	x	x		Purchase	**(+)** if consumers are in a more advanced stage of the purchase decision process and thus preferences are more specific.
	Contextual targeting	Ad campaigns on websites	[84]	x			Purchase intention	**(+)** if not combined with obtrusive advertising (e.g., pop-ups, audio content).
Trust	Retargeting (Personalized ads)	Online shopping	[98]	x	x	x	CTR intention	**(+)** if trusted retailers use personalized advertising which closely and exhaustively reflects users' interest. **(-)** if less trusted retailers use banners with high personalization depth, because of increased reactance and privacy concerns of users.
Type of collected data	Retargeting (Personalized ads)	Online banking; telecom-munications	[92]		x		Purchase intention	**(+)** if personal identification and transaction data are used to create ads that fit users' current needs. **(-)** if the high fit to users' needs evokes strong feelings of intrusiveness.
Justification	Behavioral targeting	News and query websites	[100]	x	x	x	CTR & Opt-in	Using a reciprocity argument, highlighting the free service, is generally more effective at increasing users' acceptance than a relevance argument promising users more relevant advertising in return for data disclosure.
	Personalized e-mail advertising	Movie rental and review website	[99]		x		CTR intention	**(-)** if the use of data for personalized offers is not explicitly justified and perceived utility of service is low. With higher perceived utility, reactance is less likely to occur and ad justification loses importance.
Perceived control	Segment-based targeting (Personalized ads)	Social media	[88]	x			CTR	**(+)** if users are given (perceived) control over privacy settings (although algorithm for ad targeting and personalization unchanged).
Transparency	Behavioral targeting (Personalized ads)	Social media	[95]	x		x	CTR intention	**(+)** if users are clearly informed about data collection. **(-)** if data are collected without users' awareness because of experience of vulnerability. Negative effect can be counteracted with trust-building instruments.
	Behavioral targeting	Online shopping; social media; loyalty program	[96]	x	x		CTR & purchase intention	**(+)** if users trust platform and transparency exposes information flows that users deem acceptable. **(-)** if transparency exposes information flows that users deem unacceptable due to increased privacy concerns.
Privacy regulation	Behavioral targeting	Ad campaigns on websites	[94]	x		x	Purchase intention	Regulation that restricts advertisers' ability to collect data for targeted advertising reduces targeting effectiveness.

Note: (+) = positive effect of targeting method on effectiveness measure; (-) = negative effect of targeting method on effectiveness measure.

positive opinions are not affected by these transparency measures. Consumers' *perceived control* over their personal information is also found to have a significant effect on the effectiveness of targeted advertising: Tucker finds that the number of clicks on personalized advertisements of a social network doubled after users were given a higher perceived degree of control over their personal data, although advertisers' ability to target personalized ads were unaffected by the change [88].

Similar to the case of personalized recommendations, *trust* is recognized as an important factor to alleviate privacy concerns. In particular, Bleier and Eisenbeiss demonstrate that more trusted online services can increase click-through rates by using advertising which closely and exhaustively reflects users' interest without evoking user reactance and privacy concerns [98]. Finally, *justification of ads* can influence ad effectiveness [99]. Consumers' acceptance of targeted advertising and their willingness to disclose data increases if the provider emphasizes that a service is free of charge [100]. In order to justify targeted advertising, the argument of reciprocity appears more effective than utilitarian arguments, which aim at the improved relevance of advertisements [100].

Summary 3 (Targeted advertising): *User data can frequently be used to improve ad effectiveness (e.g., click-through-rates) of targeted advertising by presenting more interesting and relevant advertisements to consumers, which increases advertising revenues (see Table 2). Online services with more personal data can target advertisements more effectively as they can exploit information on users' preferences and their purchase decision process. However, data used for targeting may also elicit feelings of intrusiveness and privacy concerns. To mitigate negative user reactions, online services can increase transparency with regard to data collection and usage, e.g., by using informational symbols, or facilitate user control over their privacy settings. In general, users are more willing to accept targeted advertising and data collection by trusted firms.*

3.4 An Overview of Economic Benefits from Data in Digital Online Markets

In summary, data has been found to be a valuable input for service personalization, recommender systems and the display of targeted advertising in order to generate economic benefits for online CSPs in the Internet economy. Figure 1 integrates the above findings and highlights that economic benefits from data are generated along different paths and in different forms. This integrated view allows us to delineate the economic benefits into three sub-categories: (i) improved customer retention, (ii) increased revenue in the consumer market, and (iii) increased revenue on other market sides. Data may contribute to these economic benefits depending on the specific use case. Specifically, service personalization can increase user satisfaction and raises switching costs. Furthermore, personalized services and recommendations improve the fit between services or products and users' preferences, thus increasing users' willingness to pay and the number of customers. As recommender systems support the discoverability of products and cross selling, they can increase consumption and attract more sellers to a platform. Targeted advertising can improve advertising effectiveness, e.g., click-through or view-through. Finally, Figure 1 also highlights the relevant moderating factors that influence the effectiveness of service personalization, recommender systems and targeted advertising in generating economic benefits from data. Note that the illustrated sub-categories are not exhaustive with respect to big data applications in general. However, based on our literature review, we conclude that they indeed capture the main benefits of big user data in the Internet economy.

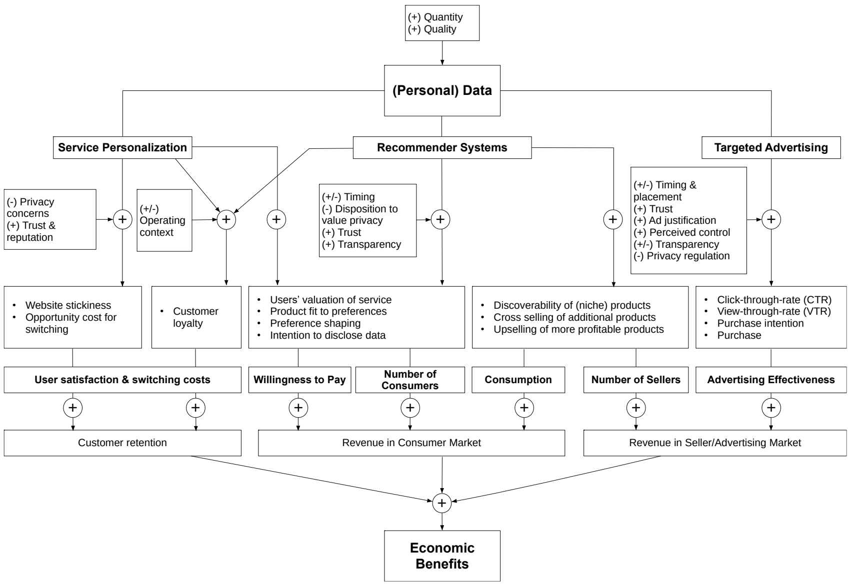

Figure 1. Overview of economic benefits from big data use in digital markets

4 Facilitating Factors for Competitive Advantages from Data

As highlighted in the previous section, (personal) data constitutes a valuable input factor for numerous digital business use cases in the Internet economy. Even if a firm's output does not directly rely on data as an input, the economic benefits from collecting and analyzing (big) data can contribute added value to products or services and increase a firm's return. In principle, however, such economic benefits can be generated by any firm. To eventually achieve a "sustained competitive advantage", as understood by the resource-based view of the firm [13, p. 102], it is necessary that a firm's data benefits cannot be imitated by its competitors [14], [101], [16]. Thus, we propose a set of facilitating factors that we derive from the nascent literature on data-driven competitive advantages in digital markets and which we summarize in Proposition 1.

Exclusive access to data: If a firm has exclusive access to larger data sets, data of higher quality, or more recent data, economic benefits are likely to translate into a long-term competitive advantage, because other firms lack the input resources to generate equivalent outputs. With regard to the replicability of personal data, there are two opposing lines of arguments: On the one hand, scholars have emphasized that data is non-rivalrous, because the same data can be used by different firms at the same time [14], [102] and the same information from consumers may also be collected by different CSPs from different services and also by different means (e.g., by tracking or by surveys).

On the other hand, there are significant economies of scale and incumbency advantages in the creation of comprehensive user profiles. First, it is much easier and less costly for firms with existing customer relations to collect first-party data than for firms with a small customer base or entrants with no customers. Thus, significant cost asymmetries are likely to arise between firms if incumbents can generate relevant data as a by-product from interactions with an established user base. Second, firms may also collect personal data as third-party data by tracking consumer behavior outside of their own content and service ecosystem. To this end, the creation of meaningful user profiles requires (i) extensive tracking across services, websites and end devices, and (ii) the ability to link the collected data to individual users or specific groups of users. With regard to (i), empirical findings show a long-tail distribution for third-party data tracking with a high concentration at the top. In 2016, trackers of only four firms were present on more than 10% of websites on the World Wide Web. However, the most widely encountered company, Google, was active on more than 70% of websites [103]. Third, intermediaries in online markets are in a special position to observe interactions and transactions that are carried out between affiliated parties over their platform. As underlying network effects promote concentrated market structures, only few firms can be expected to occupy a similar gatekeeper position. Therefore, it will be difficult for competitors to adequately replicate the scope of such data access. Finally, firms that are integrated across multiple layers of the digital value chain may leverage their position to limit the data access of potential competitors. For example, Google has announced changes to its web browser Chrome, which are expected to make it more difficult for other firms to collect third-party data, while Google can rely on first-party data [104].

Exploitative access to data: Instead of excluding competitors from data access, a data-rich firm may leverage its (exclusive) control over a particular data source by granting access to competitors and downstream services on discretionary terms. By exploiting its superior access to data through deliberate data sharing agreements, a firm may be able to generate additional economic benefits and at the same time protect or even extend its competitive advantage. From a competition perspective, this raises concerns about potential exploitative abuses, either due to the imposition of excessive and discriminatory terms and conditions [105] or due to the excessive collection of user and usage data [106]. Exploiting rather than restricting data access is appealing, because this can generate additional economic surplus, which can then be appropriated by the firm that is in control of the data resources. Such exploitation can also be implemented in the form of reciprocal data sharing agreements with other firms, e.g., through social logins as offered by Facebook, Google and Amazon [107].

Economies of scale in data analytics: Even if competitors can replicate a firm’s access to data or can find alternative input data sets, a firm may uphold a competitive advantage, because it can process data more efficiently. In particular, supply-side economies of scale are well-known to give larger firms a competitive advantage over smaller firms. Technically speaking, economies of scale occur when the average costs decline with larger output quantity, i.e., when expanding the output, the return grows more than proportionally relative to costs. In the context of big data analytics, scale economies have most prominently been discussed with regard to online search engines. Here, rare search terms ("long-tail search queries"), which increase with the total number of queries, contribute a relatively large added value to a better quality of search results, because they likely contain new information [108]. Empirical studies confirm that in many predictive analytics applications of big data, (i) there are benefits from larger data sets, (ii) these benefits are marginally decreasing as data sets become very large, and (iii) there is a minimum required scale. More precisely, de Fortuny et al. and Martens et al. demonstrate that prediction accuracy increases for larger data sets of fine-grained user behavior data [19-20]. Whereas benefits decrease marginally as prediction accuracy approaches the theoretical benchmark [109], the studies show this convergence is not yet reached in many popular application settings. For the online advertising industry, Lewis and Rao find that only for very large data sets, firms are able to reliably measure whether advertising campaigns are indeed effective [110].

Network effects and platform business models: Network effects are well-known to influence the competitive dynamics in digital markets and to manifest competitive advantages of dominant firms, due to their 'winner-takes-all' characteristic. Next to the direct impact of network effects on competition, additional issues emerge if firms generate economic benefits from big data that they collect from their users. Most notably, network effects can protect a firm's exclusive access to user data, as users are less likely to visit other businesses and share their data with these firms.

Moreover, additional self-reinforcing *feedback loops* due to data collection can create a competitive advantage. Specifically, two conceptual feedback loops are frequently conjectured [111]: First, as more users generate more data, firms can reap larger economic benefits and generate more value-added for consumers (e.g., through better service quality). In turn, this is assumed to attract more users (user feedback loop). For example, a search engine's collection and analysis of query logs is a main input for the improvement of the search results ranking, which then determines the value of the service for users [108]. Second, more data in consequence of more users also enables more effective targeted advertising and thus generates larger advertising revenues. In turn, this allows the firm to invest more in service quality or other added value for consumers. Again, this is assumed to result in more users (monetization feedback loop). Given these feedback effects, economic benefits from data may translate directly into competitive advantages as they tend to reinforce small relative advantages. First mover advantages may then quickly become sustained competitive advantages, because competitors are unable to initiate the same feedback loop.

With regard to *multi-sided platform models*, a firm's superior access to data may likewise be protected by indirect network effects. In addition, such intermediaries are able to collect additional data from interactions and transactions that are carried out between the different market sides over their platform. A digital platform is regularly able to observe the behavior and actions of the market participants on its platform [112]. In particular, the platform can collect detailed information about the demand and supply side as well as the specifics of any transaction. If the platform is in competition with some of the affiliated parties in downstream markets, as, e.g., Amazon competes with Marketplace resellers, this puts the intermediary at a direct relative competitive advantage [113]. The scale of data access that is achieved by the platform can then hardly be replicated by competitors that must share their own data with the platform.

Data-induced switching costs: The role of personal data for generating economic benefits in online markets, especially for the personalization of services and recommendations, implies a new type of data-induced switching costs [114]. These occur if switching to a competing service requires the input of data that has already been created, but the data cannot effortlessly be transferred from the current to the new service. In this case, the consumer has to incur transaction or opportunity costs in order to recreate the data at the new service. For example, switching to a competing social network requires user information, interests, photo galleries and contacts to be re-entered or re-uploaded. Moreover, some data, e.g., data observed and inferred by the CSP, may not be transferable, because the user lacks access. These data-induced costs can thus limit the freedom of choice and impede consumer switching. Hence, they may establish competitive advantages for incumbent firms even when competitors are active. Moreover, in the presence of data-induced switching costs, an incumbent has an incentive to collect the maximum amount of personal data that is still accepted by consumers if it anticipates the entry of a competitor to lock-in consumers [114].

Digital services ecosystems and economies of scope: Data as an input resource is regularly associated with economies of scope or synergies, which imply that a single company can produce goods at lower costs than if multiple firms produce these goods separately. The analysis of composite data sets, which contain data from multiple sources (e.g., usage data from various service domains) can generate additional economic benefits in comparison to isolated analyses of the individual data sets [115]. Therefore, integration across market boundaries can allow firms to provide services more efficiently than independent providers. Economies of scope may thus be seen as one driving force behind online firms' aspiration to operate not only in a single market, but to establish integrated services ecosystems across layers of the Internet value chain (e.g. providing an operating system, a search platform, applications and content as in the case of Google). Scale advantages of integrated ecosystem providers may then raise entry barriers in individual markets and promote firm concentration across markets.

Economies of scope with respect to data inputs have two main implications for firms' expansion strategies in digital markets. On the one hand, firms may have an incentive to establish integrated services and content ecosystems and to enter new markets in cases that cannot be explained by traditional theories of market power leveraging [116]. In particular, integrated ecosystems allow these firms to track users across a variety of services and obtain access to complementary user data, which is likely to improve personalization and targeting. On the other hand, firms which already generate economic benefits from data in their home market are likely to have an advantage over independent firms when they enter other existing or new markets. In this vein, data-driven network effects as characterized by Prüfer and Schottmüller can enable a firm with more data to innovate at lower costs in "connected markets", which facilitates entry into these markets and the expansion of services ecosystems [117]. Given these characteristics, firms can leverage market power from their home market and monopolize connected markets. The magnitude of scope advantages from data has particularly important ramifications for markets that are currently in the process of digitization. In these markets (e.g., the automotive industry) firms that have generated economic benefits based on physical assets in the past often compete with entrants that generate economic benefits in their home markets based on data assets [118].

Proposition 1 (Facilitating factors): *(i) Exclusive access to data, (ii) exploitative access to data, (iii) economies of scale in data analytics, (iv) network effects and platform business models, (v) data-induced switching costs, and (vi) digital services ecosystems and economies of scope can protect a firm's data-driven competitive advantage from imitation by its competitors.*

5 Conclusion

The proposed set of facilitating factors (see Proposition 1) demonstrates that a firm's data-driven competitive advantage may not necessarily be transitory. Instead, structural conditions or firms' strategies may lead to market outcomes where a dominant firm can establish long-term market power based on its collection and use of (big) data. Although from a competition policy perspective, any final assessment will require a case-by-case analysis, the existing literature provides preliminary evidence on the presence and magnitude of some of these factors, such as scale advantages from data analytics. The presented conceptual arguments and empirical findings should inform and guide policy makers and regulators when assessing competition issues in digital markets. Likewise, managers in these markets need to assess the competitive environment of their firms with respect to these factors and adjust their data-sourcing strategies accordingly.

As big data emerges as a new source for competitive advantages, this also introduces a new lever for policy interventions if authorities aim to alleviate market concentration or mitigate market power of dominant firms. In this vein, (i) transparency obligations, (ii) a right to data portability, (iii) open data access, (iv) data silos and (v) structural separation represent a set of escalating policy measures that could be building blocks of a larger policy framework to address competitive concerns in data-driven digital markets. The economic analysis and technical feasibility of these policy instruments in the digital economy present promising avenues for future research in information systems. To this end, robust assessments of the technical implementation, the design of regulatory governance models, and the economic incentives are needed. Technical design proposals (e.g., for information systems that can manage big data access transactions), theoretical analyses (e.g., on the economic incentives of data portability) as well as empirical investigations are of high and topical relevance (e.g., on the magnitude of data-driven network effects). By doing so, further research can also address limitations of this paper. Whereas the proposed facilitating factors are conceptually well-understood, empirical evidence is still scarce. In addition, big data may have further implications for competition in digital markets beyond consumer-facing online services (e.g., in the Internet of Things). Finally, big data also raises concerns about "surveillance capitalism" [119], which warrant future investigations of individual and societal risks that stem from commercial big data use.

Acknowledgements

This project was funded by the Bavarian State Ministry of Science and the Arts and coordinated by the Bavarian Research Institute for Digital Transformation (bidt).

References

1. The Economist: Data is giving rise to a new economy, https://www.economist.com/ briefing/2017/05/06/data-is-giving-rise-to-a-new-economy (Accessed: 27.08.2020).
2. European Commission: Online Platforms and the Digital Single Market Opportunities and Challenges for Europe (2016), https://ec.europa.eu/ transparency/regdoc/rep/1/2016/EN/1-2016-288-EN-F1-1.PDF (Accessed: 27.08.2020)
3. Van Reenen, J., Patterson, C.: Research: The rise of superstar firms has been better for investors than for employees (2017), https: //hbr.org/2017/05/research-the-rise-of-superstar-firms-has-been-better- for-investors-than-for-employees (Accessed: 27.08.2020)
4. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26(2), xiii-xxiii (2002)
5. Schwarz, A., Mehta, M., Johnson, N., Chin, W.W.: Understanding Frameworks and Reviews: A Commentary to Assist Us in Moving Our Field Forward by Analyzing Our Past. ACM SIGMIS Database 38(3), 29-50 (2007)
6. Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., Cleven, A.: Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. Communications of the Association for Information Systems 37(1), 205-224 (2015)
7. Shmueli, G., Koppius, O.: Predictive Analytics in Information Systems Research. MIS Quarterly 35(3), 553-572 (2011)
8. Lepenioti, K., Bousdekis, A., Apostolou, D., Mentzas, G.: Prescriptive analytics: Literature review and research challenges. International Journal of Information Management 50, 57-70 (2020)
9. Davenport, T.H., Barth, P., Bean, R.: How 'big data' is different. MIT Sloan Management Review 54(1), 22–24 (2012)
10. Müller, O., Fay, M., vom Brocke, J.: The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. Journal of Management Information Systems 35(2), 488-509 (2018)
11. Tambe, P.: Big data investment, skills, and firm value. Management Science 60(6), 1452-1469 (2014)
12. Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J.F., Dubey, R., Childe, S.J.: Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research 70, 356-365 (2017)
13. Barney, J.B.: Firm resources and sustained competitive advantage. Journal of Management 17(1), 99–120 (1991)
14. Lambrecht, A., Tucker, C.E.: Can big data protect a firm from competition? Competition Policy International, 1–8 (2017)
15. Bhatt, G.D., Grover, V.: Types of information technology capabilities and their role in competitive advantage: An empirical study. Journal of Management Information Systems 22(2), 253-277 (2005)
16. Nevo, S., Wade, M.R.: The Formation and Value of IT-Enabled Resources: Antecedents and Consequences of Synergistic Relationships. MIS Quarterly 34(1), 163-183 (2010)

17. Acquisti, A., Taylor, C., Wagman, L.: The economics of privacy. Journal of Economic Literature 54(2), 442–492 (2016)
18. Constantiou, I.D., Kallinikos, J.: New games, new rules: big data and the changing context of strategy. Journal of Information Technology 30(1), 44–57 (2015)
19. de Fortuny, E.J., Martens, D., Provost, F.: Predictive modeling with big data: is bigger really better? Big Data 1(4), 215–226 (2013)
20. Martens, D., Provost, F., Clark, J., de Fortuny, E.J.: Mining massive fine-grained behavior data to improve predictive analytics. MIS Quarterly 40(4), 869–888 (2016)
21. Aaltonen, A., Tempini, N.: Everything counts in large amounts: A critical realist case study on data-based production. Journal of Information Technology 29(1), 97–110 (2014)
22. Gandomi, A., Haider, M.: Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management 35(2), 137–144 (2015)
23. Halevy, A., Norvig, P., Pereira, F.: The unreasonable effectiveness of data. IEEE Intelligent Systems 24(2), 8–12 (2009)
24. Bourreau, M., de Streel, A., Graef, I.: Big Data and Competition Policy: Market power, personalised pricing and advertising (2017), http://www.cerre.eu/publications/big-data-and-competition-policy (Accessed: 27.08.2020)
25. Autorité de la concurrence and Bundeskartellamt: Competition Law and Data (2016), https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf;jsessionid=2ED7A555CCABEF4215FE6FE931C2294B.1_cid378?__blob=publicationFile&v=2 (Accessed: 27.08.2020)
26. Bucklin, R.E., Sismeiro, C.: Click here for internet insight: Advances in clickstream data analysis in marketing. Journal of Interactive Marketing 23(1), 35–48 (2009)
27. Pancras, J., Sudhir, K.: Optimal marketing strategies for a customer data intermediary. Journal of Marketing Research 44(4), 560–578 (2007)
28. Federal Trade Commission: Data brokers. A call for transparency and accountability (2014), https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (Accessed: 27.08.2020)
29. Sismeiro, C., Bucklin, R.E.: Modeling purchase behavior at an e-commerce web site: A task-completion approach. Journal of Marketing Research 41(3), 306–323 (2004)
30. Trusov, M., Ma, L., Jamal, Z.: Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting. Marketing Science 35(3), 405–426 (2016)
31. Yao, S., Mela, C.F.: A dynamic model of sponsored search advertising. Marketing Science 30(3), 447–468 (2011)
32. Du, R.Y., Kamakura, W.A.: Quantitative trendspotting. Journal of Marketing Research 49(4), 514–536 (2012)
33. Monopolkommission: Competition policy: The challenge of digital markets. Special Report 68 (2015), https://www.monopolkommission.de/images/PDF/SG/s68_fulltext_eng.pdf (Accessed: 27.08.2020)
34. Montgomery, A.L., Smith, M.D.: Prospects for personalization on the internet. Journal of Interactive Marketing 23(2), 130–137 (2009)
35. Tam, K.Y., Ho, S.Y.: Understanding the impact of web personalization on user information processing and decision outcomes. MIS Quarterly 30(4), 865–890 (2006)
36. Moe, W.W.: Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. Journal of Consumer Psychology 13(1-2), 29–39 (2003)
37. Chellappa, R.K., Sin, R.G.: Personalization versus privacy: An empirical examination of the online consumer's dilemma. Information Technology and Management 6(2-3), 181–202 (2005)

38. Bello-Orgaz, G., Jung, J.J., Camacho, D.: Social big data: Recent achievements and new challenges. Information Fusion 28, 45-59 (2016)
39. Saldanha, T.J., Sahaym, A., Mithas, S., Andrade-Rojas, M.G., Kathuria, A., Lee, H.H.: Turning Liabilities of Global Operations into Assets: IT-enabled Social Integration Capacity and Exploratory Innovation. Information Systems Research 31(2), 361-382 (2020)
40. Ansari, A., Mela, C.F.: E-customization. Journal of Marketing Research 40(2), 131–145 (2003)
41. Peppers, D., Rogers, M., Dorf, B.: Is your company ready for one-to-one marketing. Harvard Business Review 77(1), 151–160 (1999)
42. Bakos, J.Y.: Reducing buyer search costs: Implications for electronic marketplaces. Management Science 43(12), 1676–1692 (1997)
43. Kovacevich, A.: Google's approach to competition, https://publicpolicy.googleblog.com/2009/05/googles-approach-to-competition.html (Accessed: 27.08.2020)
44. Benlian, A.: Web personalization cues and their differential effects on user assessments of website value. Journal of Management Information Systems 32(1), 225–260 (2015)
45. Thirumalai, S., Sinha, K.K.: To personalize or not to personalize online purchase interactions: Implications of self-selection by retailers. Information Systems Research 24(3), 683–708 (2013)
46. Zhang, J.: The perils of behavior-based personalization. Marketing Science 30(1), 170– 186 (2011)
47. Malhotra, N.K., Kim, S.S., Agarwal, J.: Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. Information Systems Research 15(4), 336–355 (2004)
48. Xu, H., Dinev, T., Smith, J., Hart, P.: Information privacy concerns: Linking individual perceptions with institutional privacy assurances. Journal of the Association for Information Systems 1(12), 2798–824 (2011)
49. Adomavicius, G., Tuzhilin, A.: Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 17(6), 734–749 (2005)
50. Ansari, A., Essegaier, S., Kohli, R.: Internet recommendation systems. Journal of Marketing Research 37(3), 363–375 (2000)
51. Brynjolfsson, E., Hu, Y.J., Smith, M.D.: From niches to riches: Anatomy of the long tail. MIT Sloan Management Review 47(4), 67–71 (2006)
52. Brynjolfsson, E., Hu, Y., Smith, M.D.: Research commentary - long tails vs. super-stars: The effect of information technology on product variety and sales concentration patterns. Information Systems Research 21(4), 736–747 (2010)
53. Fleder, D., Hosanagar, K.: Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. Management Science 55(5), 697–712 (2009)
54. Oestreicher-Singer, G., Sundararajan, A.: Recommendation networks and the long tail of electronic commerce. MIS Quarterly 36(1), 65–83 (2012)
55. Schafer, J.B., Konstan, J.A., Riedl, J.: E-commerce recommendation applications. Applications of Data Mining to Electronic Commerce 5, 115–153 (2001)
56. Hinz, O., Eckert, J.: The impact of search and recommendation systems on sales in electronic commerce. Business & Information Systems Engineering 2(2), 67–77 (2010)
57. Adomavicius, G., Bockstedt, J.C., Curley, S.P., Zhang, J.: Do recommender systems manipulate consumer preferences? A study of anchoring effects. Information Systems Research 24(4), 956–975 (2013)

58. Adomavicius, G., Bockstedt, J.C., Curley, S.P., Zhang, J.: Effects of online recommendations on consumers' willingness to pay. Information Systems Research 29(1), 84–102 (2017)
59. Ghoshal, A., Kumar, S., Mookerjee, V.: Impact of recommender system on competition between personalizing and non-personalizing firms. Journal of Management Information Systems 31(4), 243–277 (2015)
60. Ekstrand, M.D., Riedl, J.T., Konstan, J.A., et al.: Collaborative filtering recommender systems. Foundations and Trends in Human–Computer Interaction 4(2), 81–173 (2011)
61. Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold- start recommendations. In: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 253–260. ACM (2002)
62. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, pp. 291–324. Springer (2007)
63. Adomavicius, G., Tuzhilin, A.: Context-aware recommender systems. In: Ricci, F., Rokach, L., Shapira, B. (eds.) Recommender Systems Handbook, vol. 2, pp. 191–226. Springer US, New York, USA (2015)
64. Organisation for Economic Cooperation and Development: Data-driven innovation: Big data for growth and well-being (2015), https://www.oecd.org/sti/ data-driven-innovation-9789264229358-en.htm (Accessed: 27.08.2020)
65. Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A.: User profiles for personalized information access. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, pp. 54–89. Springer, Berlin, Heidelberg, Germany (2007)
66. Pipino, L.L., Lee, Y.W., Wang, R.Y.: Data quality assessment. Communications of the ACM 45(4), 211–218 (2002)
67. Wand, Y., Wang, R.Y.: Anchoring data quality dimensions in ontological foundations. Communications of the ACM 39(11), 86–95 (1996)
68. O'Mahony, M.P., Hurley, N.J., Silvestre, G.C.M.: Detecting noise in recommender system databases. In: Proceedings of the 11th International Conference on Intelligent User Interfaces. pp. 109–115. ACM (2006)
69. Wang, R.Y., Strong, D.M.: Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems 12(4), 5–33 (1996)
70. Amatriain, X., Pujol, J.M., Tintarev, N., Oliver, N.: Rate it again: Increasing recommendation accuracy by user re-rating. In: Proceedings of the Third ACM Conference on Recommender Systems. pp. 173–180. ACM (2009)
71. Ho, S.Y., Bodoff, D., Tam, K.Y.: Timing of adaptive web personalization and its effects on online consumer behavior. Information Systems Research 22(3), 660–679 (2011)
72. Johar, M., Mookerjee, V., Sarkar, S.: Selling vs. profiling: Optimizing the offer set in web-based personalization. Information Systems Research 25(2), 285–306 (2014)
73. Wang, W., Benbasat, I.: Trust in and adoption of online recommendation agents. Journal of the Association for Information Systems 6(3), 72–100 (2005)
74. Wang, W., Benbasat, I.: Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. Journal of Management Information Systems 23(4), 217–246 (2007)
75. Wang, W., Xu, J., Wang, M.: Effects of recommendation neutrality and sponsorship disclosure on trust vs. distrust in online recommendation agents: Moderating role of explanations for organic recommendations. Management Science 46(11), 4967–5460 (2018)

76. Sinha, R., Swearingen, K.: The role of transparency in recommender systems. In: CHI '02 Extended Abstracts on Human Factors in Computing Systems. pp. 830–831. ACM (2002).
77. Herlocker, J.L., Konstan, J.A., Riedl, J.: Explaining collaborative filtering recommendations. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work. pp. 241–250. ACM (2000)
78. Karwatzki, S., Dytynko, O., Trenz, M., Veit, D.: Beyond the personalization–privacy paradox: Privacy valuation, transparency features, and service personalization. Journal of Management Information Systems 34(2), 369–400 (2017)
79. Awad, N.F., Krishnan, M.S.: The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Quarterly 30(1), 13–28 (2006)
80. Bagwell, K.: The economic analysis of advertising. In: Armstrong, M., Porter, R.H. (eds.) Handbook of Industrial Organization, vol. 3, pp. 1701–1844. North-Holland, Amsterdam, Netherlands (2007)
81. Evans, D.S.: The online advertising industry: Economics, evolution, and privacy. The Journal of Economic Perspectives 23(3), 37–60 (2009)
82. Tucker, C.E.: The economics of advertising and privacy. International Journal of Industrial Organization 30(3), 326–329 (2012)
83. Goldfarb, A.: What is different about online advertising? Review of Industrial Organization 44(2), 115–129 (2014)
84. Goldfarb, A., Tucker, C.: Online display advertising: Targeting and obtrusiveness. Marketing Science 30(3), 389–404 (2011)
85. Bleier, A., Eisenbeiss, M.: Personalized online advertising effectiveness: The interplay of what, when, and where. Marketing Science 34(5), 669–688 (2015)
86. Provost, F., Dalessandro, B., Hook, R., Zhang, X., Murray, A.: Audience selection for online brand advertising: privacy-friendly social network targeting. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 707–716. ACM (2009)
87. Urban, G.L., Liberali, G., MacDonald, E., Bordley, R., Hauser, J.R.: Morphing banner advertising. Marketing Science 33(1), 27–46 (2014)
88. Tucker, C.E.: Social networks, personalized advertising, and privacy controls. Journal of Marketing Research 51(5), 546–562 (2014)
89. Kox, H.L., Straathof, B., Zwart, G.: Targeted advertising, platform competition and privacy. Journal of Economics & Management Strategy 26(3), 557–570 (2017)
90. Johnson, G.A., Lewis, R.A., Nubbemeyer, E.I.: Ghost ads: Improving the economics of measuring online ad effectiveness. Journal of Marketing Research 54(6), 867–884 (2017)
91. Lambrecht, A., Tucker, C.: When does retargeting work? Information specificity in online advertising. Journal of Marketing Research 50(5), 561–576 (2013)
92. Van Doorn, J., Hoekstra, J.C.: Customization of online advertising: The role of intrusiveness. Marketing Letters 24(4), 339–351 (2013)
93. Sahni, N.S., Narayanan, S., Kalyanam, K.: An experimental investigation of the effects of retargeted advertising: The role of frequency and timing. Journal of Marketing Research 56(3), 401–418 (2019)
94. Goldfarb, A., Tucker, C.E.: Privacy regulation and online advertising. Management Science 57(1), 57–71 (2011)
95. Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., Wetzels, M.: Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. Journal of Retailing 91(1), 34–49 (2015)

96. Kim, T., Barasz, K., John, L.K., Morales, A.: Why am i seeing this ad? The effect of ad transparency on ad effectiveness. Journal of Consumer Research 45(5), 906–932 (2019)
97. Samat, S., Acquisti, A., Babcock, L.: Raise the curtains: The effect of awareness about targeting on consumer attitudes and purchase intentions. In: Proceedings of the Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017). pp. 299–318. USENIX (2017)
98. Bleier, A., Eisenbeiss, M.: The importance of trust for personalized online advertising. Journal of Retailing 91(3), 390–409 (2015)
99. White, T.B., Zahay, D.L., Thorbjørnsen, H., Shavitt, S.: Getting too personal: Reactance to highly personalized email solicitations. Marketing Letters 19(1), 39–50 (2008)
100. Schumann, J.H., von Wangenheim, F., Groene, N.: Targeted online advertising: Using reciprocity appeals to increase acceptance among users of free web services. Journal of Marketing 78(1), 59–75 (2014)
101. Wade, M., and Hulland, J. 2004. "Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research," MIS Quarterly (28:1), pp. 107-142.
102. Tucker, C.: Digital data, platforms and the usual [antitrust] suspects: Network effects, switching costs, essential facility. Review of Industrial Organization 54(4), 683–694 (2019)
103. Englehardt, S., Narayanan, A.: Online tracking: A 1-million-site measurement and analysis. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 1388–1401. ACM (2016)
104. Johnson, L.: `That would look like 70% of traffic is gone': Here are the winners and losers of Google's plans to add privacy tools to chrome (2019), https://www.businessinsider.com/winners-and-losers-google-cutting-back-on-third-party-data-2019-5?IR=T (Accessed: 27.08.2020)
105. Botta, M., Wiedemann, K.: EU competition law enforcement vis-à-vis exploitative conducts in the data economy - Exploring the terra incognita (2018), https://ssrn.com/abstract=3184119 (Accessed: 27.08.2020)
106. Bundeskartellamt: Bundeskartellamt prohibits Facebook from combining user data from different sources, https://www.bundeskartellamt.de/ SharedDocs/Meldung/EN/ Pressemitteilungen/2019/07_02_2019_Facebook.html (Accessed: 27.08.2020)
107. Krämer, J., Schnurr, D., Wohlfarth, M.: Winners, losers, and facebook: the role of social logins in the online advertising ecosystem. Management Science 65(4), 1678–1699 (2019)
108. Argenton, C., Prüfer, J.: Search engine competition with network externalities. Journal of Competition Law and Economics 8(1), 73–105 (2012)
109. Li, X., Ling, C.X., Wang, H.: The convergence behavior of naive bayes on large sparse datasets. ACM Transactions on Knowledge Discovery from Data (TKDD) 11(1), 10:1–10:24 (2016)
110. Lewis, R.A., Rao, J.M.: The unfavorable economics of measuring the returns to advertising. The Quarterly Journal of Economics 130(4), 1941–1973 (2015)
111. Lerner, A.V.: The role of 'big data' in online platform competition (2014), https://ssrn.com/abstract=2482780 (Accessed: 27.08.2020)
112. Hagiu, A., Teh, T.H., Wright, J.: Should Amazon be allowed to sell on its own marketplace? (2020). https://ssrn.com/abstract=3606055 (Accessed: 27.08.2020)
113. Mattioli, D.: Amazon Scooped Up Data From Its Own Sellers to Launch Competing Products. (2020), https://www.wsj.com/articles/amazon-scooped-up-data-from-its-own-sellers-to-launch-competing-products-11587650015 (Accessed: 27.08.2020)
114. Wohlfarth, M.: Data portability on the internet. Business & Information Systems Engineering 61(5), 551-574 (2019)

115. Martens, B.: An economic policy perspective on online platforms (2016), https://ec.europa.eu/jrc/sites/jrcsh/files/JRC101501.pdf (Accessed: 27.08.2020)
116. Whinston, M.D.: Tying, foreclosure, and exclusion. The American Economic Review 80(4), 837–859 (1990)
117. Prüfer, J., Schottmüller, C.: Competing with big data (2017), https://ssrn.com/abstract=2918726 (Accessed: 27.08.2020)
118. Buntz, B.: The battle for the connected car market shifts into high gear (2016), https://www.iotworldtoday.com/2016/09/10/battle-connected-car-market-shifts-high-gear/ (Accessed: 27.08.2020)
119. Zuboff, S.: Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. Journal of Information Technology 30(1), 75-89 (2015)